崖壁植物太行菊与长裂太行菊全基因组大小及特征分析

王祎玲*，臧恩，张昊，刘志霞，兰亚飞，何珊，郝伟丽，曹艳玲

（山西师范大学 生命科学学院，太原 030031）

摘要：太行菊（*Opisthopappus taihangensis*）、长裂太行菊（*Opisthopappus longilobus*），为太行山特有多年生崖壁草本植物，菊科（Compositae）重要野生资源，具有较高的经济与生态价值。为确定适合两物种的全基因组测序策略，本研究利用流式细胞法和高通量测序技术，分析两物种基因组大小、杂合率、重复序列及 GC 含量等信息。结果表明：（1）流式细胞法估算太行菊基因组大小为 2.1 Gb，长裂太行菊基因组大小 2.4 Gb；（2）高通量测序修正后太行菊基因组大小为 3.13 Gb，重复序列比例为 84.35 %，杂合度为 0.99 %，GC 含量为 36.56 %；长裂太行菊基因组为 3.18 Gb，重复序列比例为 83.83 %，杂合度为 1.17 %，GC 含量为 36.62 %；（3）初步组装后 GC 含量分布及平均深度存在异常，出现分层现象，可能是两物种基因组杂合率较高所致。从基因组结构看，太行菊、长裂太行菊均属于高重复、高杂合、大基因组的复杂基因组，建议使用 Illumina + PacBio 测序组装策略，进行全基因组测序分析。

关键词：太行菊，长裂太行菊，基因组调查，基因组大小

中图分类号：　　**文献标识码**：A　　文章编号：

Genome sizes and characteristics of cliff plants *Opisthopappus taihangensis* and *O. longilobus* on the Taihang Mountains

WANG Yiling*, ZANG En, ZHANG Hao, LIU Zhixia, LAN Yafei,

HE Shan, HAO Weili, CAO Yanling

(*College of Life Science, Shanxi Normal University*, Taiyuan 030031, China)

Abstract: *Opisthopappus taihangensis* and *O. longilobus*, being perennial cliff herbs and endemic on the Taihang Mountains, an important wild germplasm resource of Compositae, have high economic and ecological values. To determine the appropriate sequencing strategy for the whole genome of *O. taihangensis* and *O. longilobus*, in this study, the genome sizes, heterozygosity, GC content, and repeatability were analyzed through the flow

基金项目：国家自然科学基金资助项目（31970358）；山西省回国留学人员科研项目（2020-090）；山西省研究生教育创新项目（2020SY323）[Supported by the National Natural Science Foundation of China (31970358); Research Project of Shanxi Scholarship Council (2020-090); Graduate Education Innovation Project of Shanxi (2020SY323)] 。

作者简介：王祎玲（1970--），博士，教授，主要从事植物分子生态学研究，（E-mail）ylwangbj@hotmail.com。

*通信作者

cytometry and high-throughput sequencing methods. The results were as follows: ⑴ Using maize of known genome size as controls, the genome size of *O. taihangensis* was approximately 2.1 Gb, while that of *O. longilobus* was approximately 2.4 Gb; ⑵ For *O. taihangensis*, the revised genome size was 3.13 Gb, and the repetitive sequences proportion, heterozygosity and GC content in the whole genome were estimated to be 84.35 %, 0.99 % and 36.56 % respectively. Within *O. longilobus*, the revised genome size, the proportion of repetitive sequences, heterozygosity and GC content were 3.18 Gb, 83.83 %, 1.17 % and 36.62 % respectively; ⑶ The initial depth and content distribution of GC appeared abnormal after initial assembly, which might be related to the relatively high heterozygous rate of the two species. Above all, the whole genome of *O. taihangensis* and *O. longilobus* both were a large and complex genome with high heterozygosity and repetitiveness. Therefore, it suggests that the use of Illumina + PacBio sequencing assembly strategy for the whole genome sequencing analysis of two *Opisthopappus* species in the future.

Key words: *Opisthopappus taihangensis*, *O. longilobus*, genome survey, genome size

随着第二代高通量测序技术（next-generation sequencing technology）和第三代单分子测序技术（single-molecule sequencing）的发展与成熟，测序的时间成本不断降低，为不同物种基因组测序提供了条件和便利(Aird et al., 2011；施季森等, 2012；Li et al., 2019)。而物种全基因组测定可为其基因组学和进化生物学领域研究提供线索参考，为其分子生物学、转录组学、生物信息学研究奠定基础。从基因组水平出发，分析植物物种的生长、适应、进化等问题，可很好促进对植物的进一步认识，同时加快新基因的发现、挖掘与利用（Bi et al., 2019；李江莹等, 2020；赵乐等, 2021）。但在大规模实施深度测序之前，有必要进行基因组 Survey，通过 Survey 判断植物物种基因组大小及复杂程度，提前了解植物基因组基本情况，减少测序盲目性，并据此选取合适测序策略和序列拼接软件（唐其等, 2015；霍恺森等, 2018；Bi et al., 2019；Li et al., 2019；郑燕等, 2020）。

太行菊（*Opisthopappus taihangensis*）与长裂太行菊（*O. longilobus*）均为太行菊属（*Opisthopappus*）多年生草本植物，仅生长在太行山的悬崖裂缝中，典型的崖壁植物，具有良好的抗旱抗寒性（Chai et al., 2018, 2020）。作为二倍体植物（2n=18），太行菊、长裂太行菊是菊亚族（Chrysantheminae）中较为原始的物种（Ye et al., 2021），菊科（Compositae）重要的野生种质资源，其体内可能含有如耐寒耐旱等方面的大量优异基因，是菊科种质创新的良好基因源。然而太行菊、长裂太行菊全基因组大小以及基因组特征等信息较为缺乏，阻碍其基因组测序工作的进一步开展，也不利于其进化生物学等相关工作的研究 （霍恺森等, 2018）)。因此，开展太行菊与长裂太行菊全基因组的测序工作十分必要，这将从分子水平揭示其适应、演化、抗性等机理，并为其经济价值的综合合理利用提供科学理论支持 (宋立肖等, 2018) 。

自模式植物拟南芥（*Arabidopsis thaliana*）的全基因组公布以来，已有 400 多种植物的基因组被测序（Chen et al., 2018, 2019; https://www.plabipd.de/index.ep），而且目前还有许多植物物种正在测序，这为研究植物全基因组测序提供了大量的参考信息，特别是同属菊科管状花亚科（Carduoideae）的青蒿（*Artemisia annua*）（Shen et al., 2018）、野菊（*Chrysanthemum nankingense*）（Song et al., 2018）、太阳花（*Helianthus annuus*）（Badouin et al., 2017）、小蓬草（*Conyza canadensis*）（Peng et al., 2014）、刺苞菜蓟（*Cynara cardunculus*）（Scaglione et al., 2016）全基因组的完成，对于太行菊、长裂太行菊基因组的解析具有重要的借鉴作用。

在系统关系上，太行菊属与菊属（*Dendranthema*）、亚菊属（*Ajania*）的大部分类群组成菊属群，其位置更接近于蒿亚族（Artemiisinae）（赵宏波等, 2010; Zhao et al., 2010）。同为蒿亚族（Oberprieler et al., 2007）的菊属植物野菊和蒿属植物青蒿，其基因组均为高重复、高杂合、大基因组的复杂基因组（Shen et al., 2018; Song et al., 2018），而与其系统关系相近的太行菊属两物种的基因组如何？是否呈现出相似的基因组特征？

为此，本研究通过流式细胞法（Arumuganathan & Earle, 1991；Doležel et al., 2007）和高通量测序技术，拟解决以下问题：⑴ 预估太行菊与长裂太行菊基因组大小；⑵ 测定、评估两物种全基因组大小和特点。研究结果旨在全面了解太行菊与长裂太行菊基因组特征，为后续其全基因组 de novo 测序及组装策略提供依据，为挖掘其抗旱抗寒基因及利用其潜在的基因资源提供线索 (霍恺森等, 2018) 。

1 材料与方法

1.1 材料

2019 年 10 月，从山西王莽岭（太行菊生长点）和壶关（长裂太行菊生长点）收集太行菊和长裂太行菊的种子并带回实验室，于 2020 年 1 月在实验室进行萌发，然后盆栽培养。6 月挑选长势良好的健康个体（每个物种 3 个个体），选取完整叶片，液氮速冻后，置于-80 ℃超低温冰箱中保存备用 （宋立肖等, 2018） 。

1.2 流式细胞法检测

各取两物种 0.5 cm 叶片，将其放在平底培养皿中，加入 400 μl 提取 OTTO 缓冲液，用刀片垂直把叶片切碎，持续 30~60 s。室温下，孵育 30~90 s，然后过滤器过滤，加入 1.6 ml 染色溶液（染色缓冲液 ＋PI＋RNase 储存液），避光室温，孵育 30~60 min。最后在 Sysmex CyFlow® Cube6 流式细胞仪上进行检测。

将已知基因组大小的玉米（大小约 2.3 Gb）作为对照样品，首先将玉米、太行菊、长裂太行菊单独测定，检测每个样品的相对荧光强度；然后以玉米为对照，对玉米与太行菊、玉米与长裂太行菊的混合样品进行相对荧光强度的检测。根据不同样品相对荧光强度峰值，参考对照样品的基因组，估算太行菊、长裂太行菊的基因组大小。

1.3 基因组 Survey 分析

1.3.1 DNA 提取

用改良 CTAB 法对太行菊、长裂太行菊叶片基因组 DNA 进行提取。随后用分光光度计和琼脂糖凝胶电泳分别对提取的 DNA 纯度、浓度和完整性进行检测 （赵乐等, 2020） 。

1.3.2 样品测序

将太行菊、长裂太行菊（各 3 个样品）的 DNA 样品委托杭州联川生物公司进行测序。通过 Covaris 超声波破碎仪，随机打断成片段，经末端修复-加 A 尾-加测序接头-纯化-PCR 扩增等完成整个文库制备。构建好的文库通过 Illumina Hiseq 进行 PE（双末端，Pair-end）测序。

测序得到的原始序列，含有低质量带接头的 reads，会对后续分析带来影响，对 raw reads 进行精细过滤，得到 clean reads，用于后续基因组大小、杂合度、GC 含量等分析。以 Q20、Q30 作为衡量测序质量的指标，当 Q20≥90 %、Q30≥80 %，表明测序数据质量较好。

1.3.3 污染评估

所测物种基因组 DNA 样品，如果存在污染，不仅会降低有效数据量，还会影响基因组 Survey 分析结果的准确性，所测物种基因组评估结果将会呈现误差，导致基因组组装策略出现偏差，影响后续基因组的组装效果 （赵乐等, 2020）。

为了判断提取的太行菊、长裂太行菊的基因组 DNA 是否受到污染，从过滤后的高质量 clean reads 数据中随机抽取 10 000 条 reads，用 Blast 软件比对 NCBI 核苷酸数据库（NT 库），如果是同源比对，则认为样本不存在外源污染。

1.3.4 基因组特征预估

为了对基因组大小有个大致判断，用 K-mer 法进行（Liu et al., 2013；Chen et al., 2015）。以每个 K-mer 深度（depth）为横坐标，K-mer 频度（frequency）为纵坐标，绘制 K-mer 深度频度分布图，根据曲线估计 K-mer 的深度值，对基因组大小进行估测。从测序数据中提取出的长度，视为 K-mer 的寡聚核苷酸序列，对太行菊、长裂太行菊预测序列有效数据进行 K=17 分析，根据公式：基因组大小 = 总碱基数 / 平均测序深度 = 总 K-mer 数 / 平均 K-mer 深度，计算基因组大小 （霍恺森等, 2018; 赵乐等, 2020）。

通过贝叶斯模型，根据 K-mer 的频率数和深度值，通过迭代修正所测物种基因组的杂合度和重复序列。根据杂合种类数百分比、纯合种类数百分比、所有种类数，计算太行菊、长裂太行菊的杂合度。同时，计算标准泊松分布及实际数据曲线峰值后的面积差，计算两物种基因组重复序列的百分比。

1.3.5 基因组组装

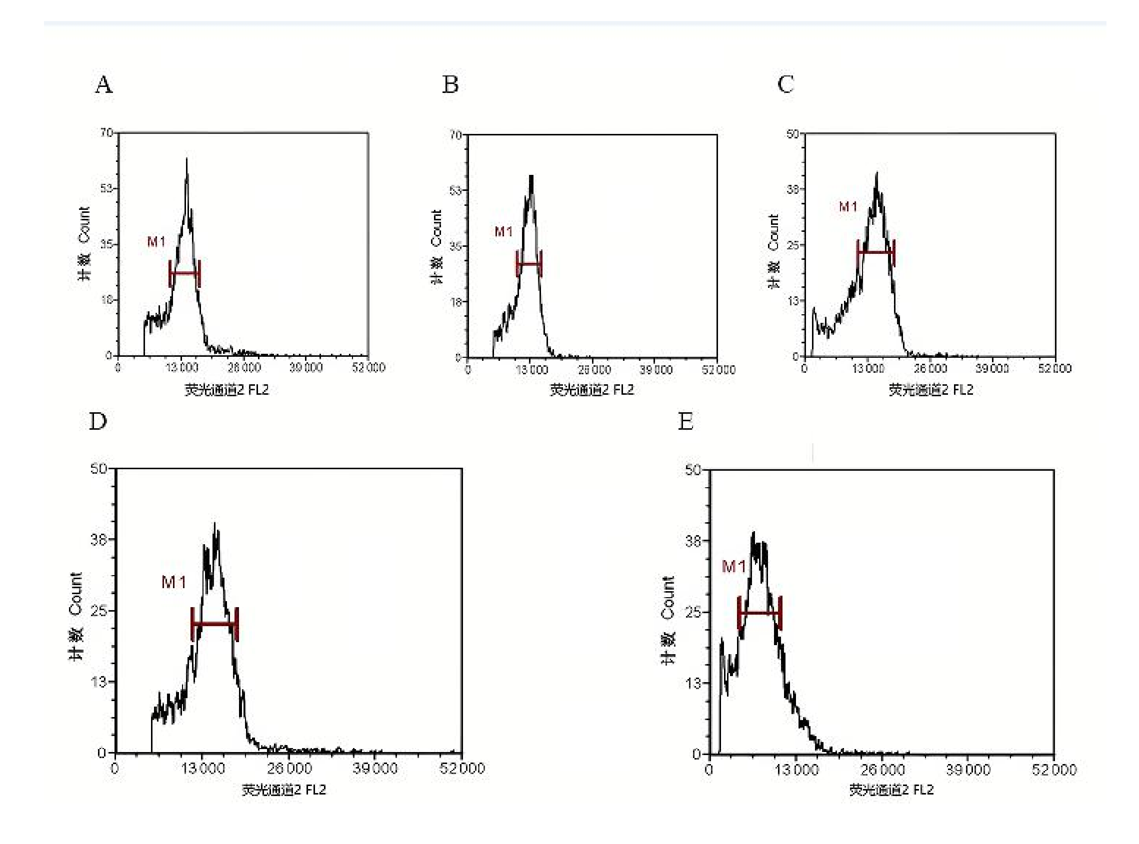
运用 Soapdenovo 软件（Vurture et al., 2017），对太行菊、长裂太行菊的 clean reads 有效序列进行拼接，用 K=41 组装到 contig 和 scaffold。将组装好的基因组序列与 raw reads 进行比对，分析组装序列的 GC 含量、contig 覆盖深度、长度和数量分布。

根据基因组测序序列 GC depth 的分布图，分析两物种测序序列是否有明显的 GC 偏向。一般高 GC 或低 GC 区域，测序深度与正常区域会存在较大差异，覆盖度较低。本研究中，以 10 kb 无重叠区域作为窗口，计算太行菊、长裂太行菊基因组的 GC 含量。

2 结果

2.1 流式细胞法估测太行菊、长裂太行菊基因组大小

对玉米、太行菊、长裂太行菊单独测定，首先检测每个样品的相对荧光强度（图 1：A-C），结果显示，玉米、太行菊、长裂太行菊样品基因组 DNA 相对荧光强度的峰值分别为 58、56、41。玉米与太行菊混合样品基因组 DNA 相对荧光强度的峰值在 40 左右（图 1：D），玉米与长裂太行菊混合样品相对荧光强度的峰值同样在 40 左右（图 1：E）。根据流式细胞结果，太行菊基因组大小估算为 2.1 Gb，长裂太行菊基因组大小约为 2.4 Gb。



M1. Marker； **A**. 玉米单独样品；**B**. 太行菊单独样品；**C**. 长裂太行菊单独样品；**D**. 太行菊与玉米混合样品；**E**. 长裂太行菊与玉米混合样品。

M1. Marker； **A**. Maize samples; **B**. *O. taihangensis* samples; **C**. *O. longilobus* samples; **D**. Mixed samples of *O. taihangensis* and maize; **E**. Mixed samples of *O. longilobus* and maize.

图 1 流式细胞法检测结果

Fig. 1 Flow cytometry of *Opisthopappus taihangensis* and *O. longilobus*

2.2 Survey 分析太行菊、长裂太行菊基因组大小

2.2.1 测序产量统计

通过文库构建，太行菊产出原始数据为 99.94 Mb，过滤后高质量数据 22.67 Mb；长裂太行菊原始数据 109.74 Mb，过滤后高质量数据 80.49 Mb。不论太行菊还是长裂太行菊，测序数据 Q20 均在 97.42 %以上，Q30 均为 92.53 %以上，测序错误率 0.04 %（正常范围为<0.05 %），表明测序质量较好，可进行进一步的后续分析。

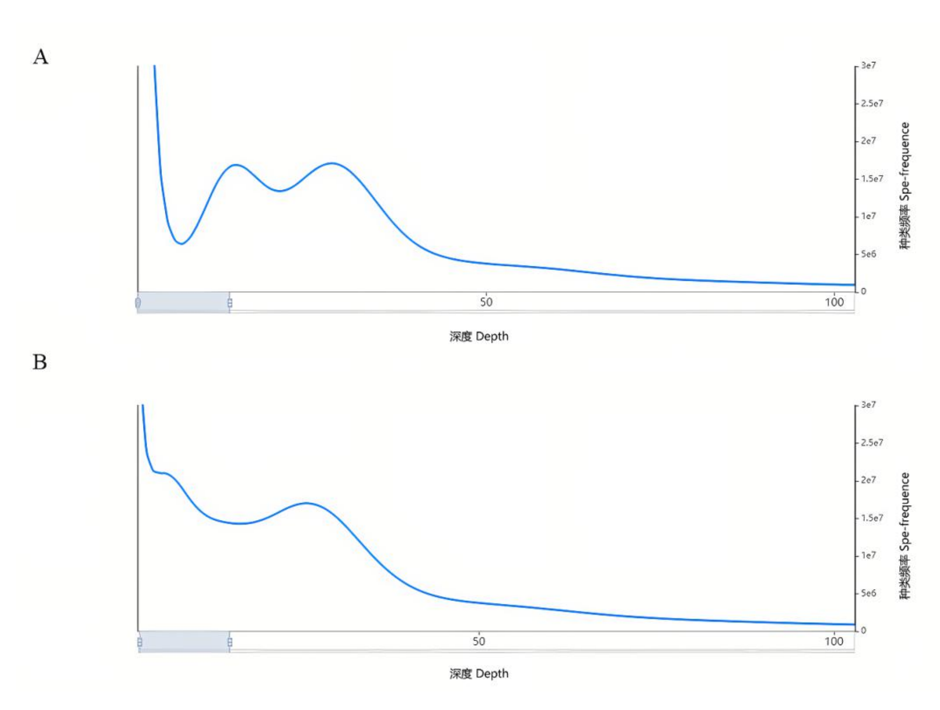
2.2.2 样品污染评估

10 000 条随机抽取的 clean reads 在 NT 库中进行同源比对，发现太行菊比对到 *Artemisia frigida*、 *Chrysanthemum indicum*、 *Helianthus maximiliani*、 *Chrysanthemum x* 分别占比对上 NT 库 reads 数的 1.56 %、0.72 %、0.54 %、0.27 %。长裂太行菊比对到上述四个物种分别占比对上 NT 库 reads 数的 1.09 %、0.48 %、0.31 %、0.09 %。*A. frigida* 、*C. indicum* 与太行菊、长裂太行菊同属于菊亚族，但与 *A. frigida* 亲缘关系更近，比对所占比例较高。由于太行菊、长裂太行菊基因组信息未知，在 NT 库中基因注释极少，所以与其它物种比对比例较低。

另外，比对结果中未发现动物、微生物等异常比对，表明太行菊、长裂太行菊基因组 DNA 样品测序数据没有污染，可用于 Survey 分析。

2.2.3 基因组大小估计

在图 2 中，太行菊、长裂太行菊的 K=17 曲线具有严重拖尾，暗示都有很高的重复序列比例。在 depth=28 附近，出现太行菊主峰值，由公式 Kmer-number / depth 计算得到太行菊基因组大小为 3.15 Gb 左右，修正后基因组大小为 3.13 Gb，基因组的杂合率为 0.99 %，重复序列比例为 84.35 %（表 1）；depth=26 时出现长裂太行菊的主峰值，基因组大小为 3.20 Gb 左右，修正后的基因组大小 3.18 Gb，基因组杂合率 1.17 %，重复序列比例 83.83 %（表 1）。由此可知，太行菊属太行菊与长裂太行菊都为高重复、杂合基因组。



A. 太行菊；**B**. 长裂太行菊。

A. *O. taihangensis*; **B**. *O. longilobus*.

图 2 太行菊、长裂太行菊的 K-mer 分布曲线

Fig. 2 K-mer distribution curve of *Opisthopappus taihangensis* and *O. longilobus*

表 1 太行菊、长裂太行菊 K-mer 分析数据统计

Table 1 K-mer analysis data of *Opisthopappus taihangensis* and *O. longilobus*

物种 Species	K-mer 总数量 K-mer total number	K-mer 深度 K-mer depth	基因组大小（Gb） Genome size	修正后基因组大小（Gb） Revised genome size	杂合率（%） Heterozygous	重复序列比例（%） Repeat sequence ratio	GC 含量（%） GC content
太行菊 *Opisthopappus taihangensis*	94473837553	30	3.15	3.13	0.99	84.35	36.56
长裂太行菊 *O. longilobus*	89639775705	28	3.20	3.18	1.17	83.83	36.63

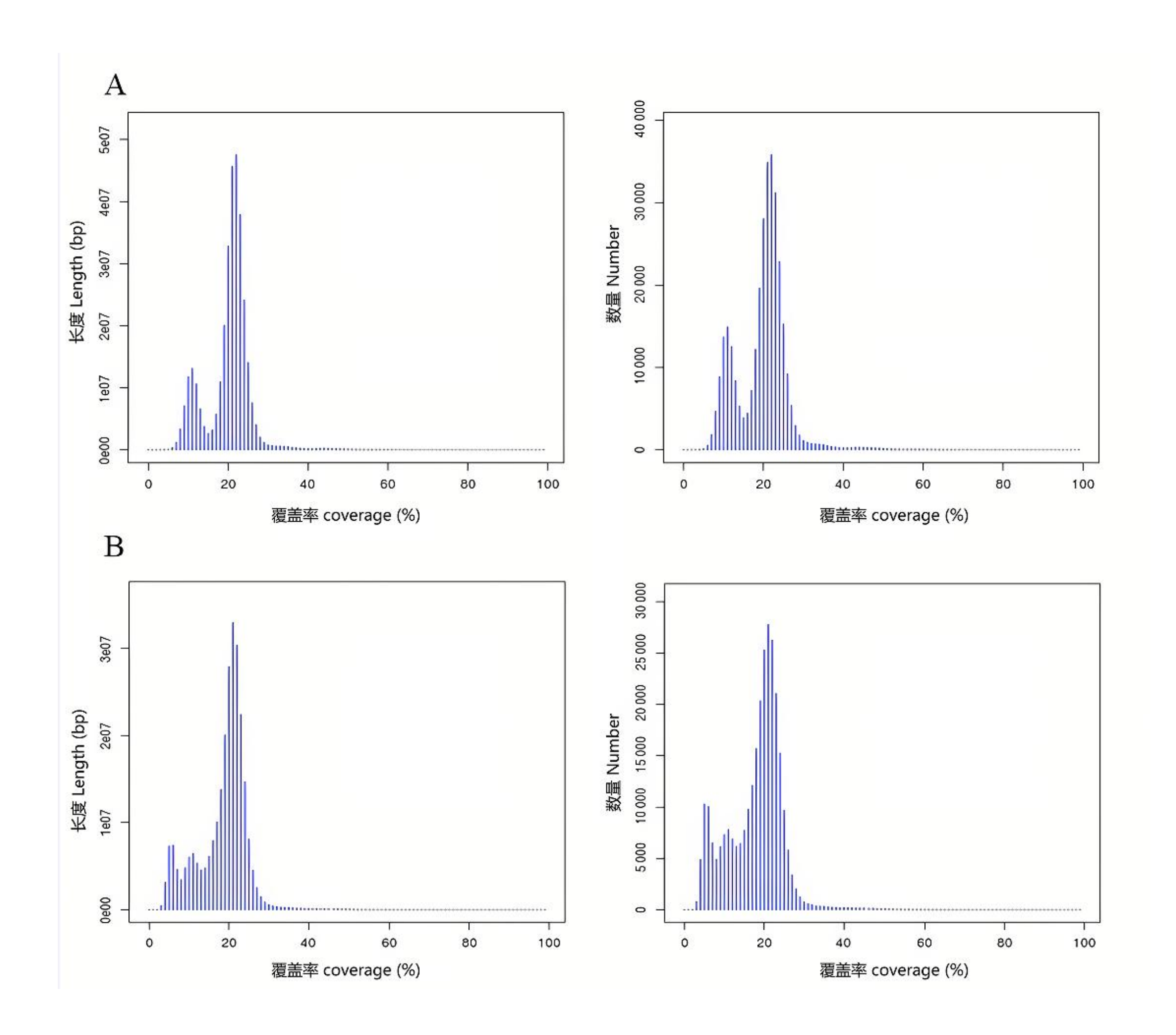
2.2.4 基因初步组装结果

在太行菊中，共得到 4 148 869 条 contigs，序列总长为 1.19 Mb，Contig N50 长度 445 bp，N90 114 bp，最长序列长度为 24 674 bp，进一步组装后得到 3 885 802 条 scaffolds，总长 1.22 Mb，最长序列长度为 24 674 bp，Scaffold N50 为 510 bp，N90 118 bp（表 2）。长裂太行菊中，共有 4 776 945 条 contigs，序列总长为 1.30 Mb，Contig N50 长度 408 bp，N90 113 bp，最长序列长度为 24 198 bp，进一步组装后 4 453 317 条 scaffolds，总长 1.34 Mb，最长序列长度为 24 198 bp，Scaffold N50 为 477 bp，N90 116 bp（表 2）。其中，Contig N50 和 Scaffold N50 的长度较短，可能是由于两个物种基因组杂合率在 0.99 %以上所致。从图 3 中，可以看出明显的峰，主峰前的峰为杂合峰，主峰后的峰为重复峰。不论是太行菊还是长裂太行菊，峰值在 20 X左右时的峰值为纯合峰，初步判断两物种基因组为复杂基因组。

表 2 太行菊、长裂太行菊基因组组装统计

Table 2 Genome assembly of *Opisthopappus taihangensis* and *O. longilobus*

物种 Species	项目 Item	Contig	Scaffold
太行菊 *Opisthopappus taihangensis*	总长度 Total length（Mb）	1.19	1.22
	总数量 Total number	4 148 869	3 885 802
	最大长度 Max length （bp）	24674	24674
	N50 长度 N50 length （bp）	445	510
	N90 长度 N90 length （bp）	114	118
长裂太行菊 *O. longilobus*	总长度 Total length （Mb）	1.30	1.34
	总数量 Total number	4 776 945	4 453 317
	最大长度 Max length （bp）	24198	24198
	N50 长度 N50 length （bp）	408	477
	N90 长度 N90 length（bp）	113	116



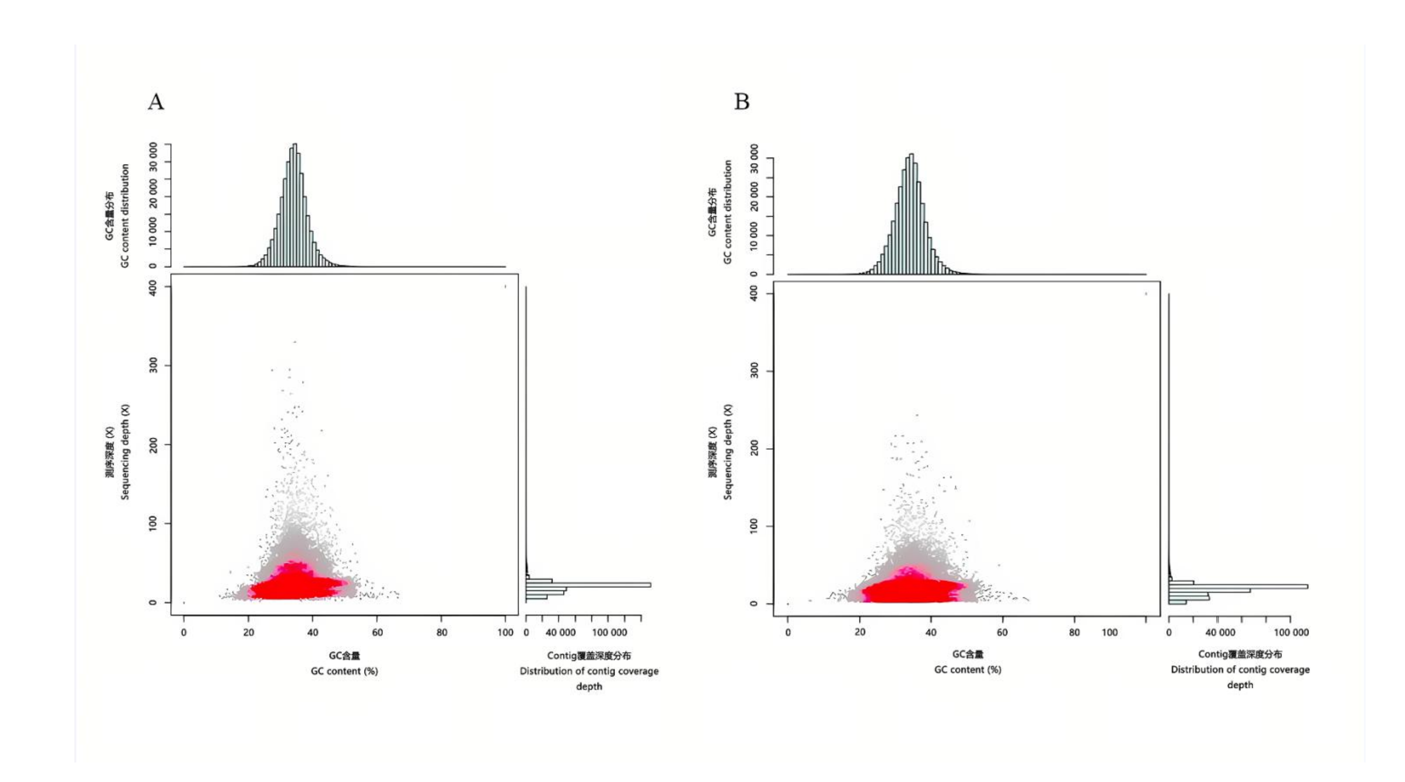
A. 太行菊 Contig 覆盖深度和长度、数量分布图；**B**. 长裂太行菊 Contig 覆盖深度和长度、数量分布图。**A**. Contig coverage depth, length and quantity distribution of *O. taihangensis*; **B**. Contig coverage depth, length and quantity distribution of *O. longilobus*.

图 3 Contig 分布图

Fig. 3 Contig distribution

2.2.5 GC 含量及分布情况

太行菊属两物种中，窗口的GC含量几乎全部处于20~60 %之间，并主要集中在30 %左右，其中太行菊 GC 含量为 36.56 %，长裂太行菊为 36.63 %（表 1）。两物种样品不存在明显异常，GC 含量没有明显偏向，GC depth 深度的分布可分为三层：高中低深度 3 个区域（图 4）。中深度区域大约为高深度区域的 50 %左右，可能与太行菊、长裂太行菊一定的杂合有关。在组装过程中，杂合可能会导致同源染色体杂合部位单条组装，引起 GC 含量出现分层现象。



A. 太行菊；**B**. 长裂太行菊。

A. *O. taihangensis*; **B**. *O. longilobus*.

图 4 Contig GC 含量和覆盖深度

Fig. 4 GC content and depth of coverage

3 讨论

本研究通过流式细胞法和高通量测序 K-mer 法初步调查了菊科崖壁植物太行菊、长裂太行菊的基因组大小、杂合率和 GC 含量等特征（宋立肖等, 2018）。太行菊、长裂太行菊基因组初步估算大小分别为2.1 Gb和2.4 Gb；修正后，基因组大小分别为3.13 Gb、3.18 Gb。

目前公布的菊科植物中，最小的基因组为管状花亚科紫菀族（Astereae）的小蓬草，其基因组只有 335 Mb（Peng et al., 2014），最大的是春黄菊族（Anthemideae）植物，基因组约为 138.88 Gb（Garcia et al., 2013）。与太行菊、长裂太行菊亲缘关系较近的青蒿、野菊，基因组大小分别为 1.74 Gb（Shen et al., 2018）和 3.07 Gb（Song et al., 2018）。太行菊属太行菊、长裂太行菊基因组大小符合菊科植物的基因组特征（Garcia et al., 2013）。物种的进化与 DNA 含量之间的关系是复杂的，基因组越大的物种，其濒危程度越高（Vinogradov, 2003）。相对于野菊、青蒿等，太行菊、长裂太行菊生长在相对恶劣的崖壁环境下，已被列为国家二级濒危物种。

本研究中，流式细胞法估测结果比 K-mer 分析法的结果小大约 1 Gb。流式细胞法估测太行菊、长裂太行菊基因组较小，可能与选择基因组较小的玉米为对照样品（2.3 Gb）有关，而 K-mer 分析基因组法是基于数学计算可能更全面准确（Doležel et al., 2007; Wang et al., 2018）。在野菊基因组分析中，流式细胞法估测的结果大于 K-mer 分析结果（Song et al., 2018），而其它植物的基因组调查，两种方法估测的结果也不完全一致，如甘薯

属的马鞍藤（*Ipomoea pes-caprae*）（霍恺森等, 2019）、绣球属的绣球（*Hydrangea macrophylla*）（陈双双等, 2021）

在已释放的植物基因组中，GC 含量大多在 30~47 %（邓果特等, 2013；宋立肖等, 2019；于福来等, 2019）。太行菊、长裂太行菊基因组 GC 含量分别为 36.56 %和 36.63 %。同属管状花亚科的刺苞菜蓟基因组 GC 含量为 32 %（Scaglione et al., 2016），青蒿中 GC 含量 31.5 %（Shen et al., 2018），野菊为 37.2 %（Song et al., 2018）。太行菊、长裂太行菊基因组 GC 含量在所释放的植物基因组 GC 含量范围之内。

根据基因组杂合度大小，基因组分为微杂合基因组（0.5 % ≤杂合率 ＜0.8 %）、高杂合基因组（杂合率 ≥ 0.8 %）和高重复基因组（重复序列比例 ≥ 50 %）（伍艳芳等, 2014；周佳熠等, 2017；王雪等, 2018）。太行菊、长裂太行菊基因组杂合率分别为 0.99 %和 1.17 %，重复序列比例分别为 84.35 %和 83.83 %。植物基因组杂合受繁殖方式影响，一般具有自花授粉繁育方式的物种的杂合度低于异花授粉植物（王雪等, 2018；都明理等, 2019）。太行菊、长裂太行菊自交不亲和（胡枭和赵恩惠, 2008），可进行有性生殖形成种子，也可在茎节处长出新的枝条，即有性繁殖和无性繁殖并存，这种繁育方式使得太行菊、长裂太行菊基因组存在一定的杂合率。青蒿基因组的杂合率在 1.0~1.5 %之间，重复序列比例 61.57 %（Shen et al., 2018）；野菊基因组也呈现出高的杂合率且重复序列为 69.6 %（Song et al., 2018）。具有高比例的重复序列是菊科已公布的大基因组物种（3 Gb）的共有特征（Garcia et al., 2013）。这些高比例的重复序列在一定程度上加大了太行菊、长裂太行菊的基因组。

综上，从基因组基本结构特征看，太行菊和长裂太行菊基因组都属于高重复、高杂合、大基因组的复杂基因组。

另外，采用 K-mer＝41 进行基因组初步组装，太行菊 Contig N50 为 445 bp，进一步组装后 Scaffold N50 为 510 bp，最长序列长度为 24 674 bp。长裂太行菊 Contig N50 长度 408 bp，进一步组装后 Scaffold N50 为 477 bp，最长序列长度为 24 198 bp。根据两物种基因组的调查分析，建议后续研究采用第二代和第三代测序技术相结合，对太行菊、长裂太行菊基因组进行测序和组装，同时，辅以 Hi-C，进行染色体水平组装，以期获得两物种高质量的全基因组图谱。

本研究获得的太行菊属太行菊、长裂太行菊基因组大小、特征等信息，为以后绘制两物种基因组的精细图谱奠定了基础，也为研究利用菊科野生种质资源提供了参考(赵乐等, 2020)。

参考文献：

AIRD D, ROSS MG, CHEN WS, et al., 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries [J]. Genome Biol, 12(2): R18.

AREMUGANATHAN K, EARLE ED, 1991. Estimation of nuclear DNA contents of plants by flow cytometry [J]. Plant Mol Biol Rep, 9(3): 229-241.

BADOUIN H, GOUZY J, GRASSA CJ, et al., 2017. The sunflower genome provides insights into oil metabolism, flowering and asterid evolution [J]. Nature, 546: 148–152.

BI QX, ZHAO Y, CUI YF, et al., 2019. Genome survey sequencing and genetic background characterization of yellow horn based on next-generation sequencing [J]. Mol Biol Rep, 46(4): 4303-4312.

CHAI M, WANG S, HE J, et al., 2018. De novo assembly and transcriptome characterization of *Opisthopappus* (Asteraceae) for population differentiation and adaption [J]. Front Genet 9:371.

CHAI M, YE H, WANG Z, et al., 2020. Genetic divergence and relationship among *Opisthopappus* species identified by development of EST-SSR markers [J]. Front Genet, 11:177.

CHEN F, DONG W, ZHANG J, et al., 2018. The sequenced angiosperm genomes and genome databases [J]. Front Plant Sci, 9: 418.

CHEN F, SONG Y, LI X, et al., 2019. Genome sequences of horticultural plants: past, present, and future [J]. Hortic Res, 6:112.

CHEN SS, LI XY, FENG J, et al., 2021. Genome size and characterization of *Hydrangea macrophylla* based on flow cytometry and genome survey [J]. Jiangsu Agr Sci, 49(12): 39-44. [陈双双, 齐香玉, 冯景, 等, 2021. 基于流式细胞术和基因组 Survey 的绣球基因组大小及特征分析 [J]. 江苏农业科学, 49(12): 39-44.]

CHEN W，HASEGAWA D，ARUMUGANATHAN K，et al. 2015. Estimation of the whitefly *bemisia tabaci* genome size based on k-mer and flow cytometric analyses [J]. Insects, 6: 704-715.

DENG GT, LIU QB, JIANG JX, et al., 2013. Estimation of genome size of *Miscanthus floridulus* [J]. J Plant Genet Resour, 14(2): 339-341. [邓果特, 刘清波, 蒋建雄, 等, 2013. 五节芒基因组大小测定 [J]. 植物遗传资源学报, 14(2): 339-341.]

DOLEŽEL J, GREIHUBER J, SUDA J. 2007. Estimation of nuclear DNA content in plants using flow cytometry [J]. Nat Protoc, 2(9): 2233-2244.

DU ML, XU J, ZHU CR, et al., 2019. Chromosome number determination and genome size estimation of three Bupleurum species [J]. Jiangsu Agr Sci, 47(11): 191-193. [都明理, 徐娇, 朱楚然, 等, 2019. 3 种柴胡染色体数目测定及基因组大小估测[J]. 江苏农业科学, 47(11): 191-193.]

GARCIA S, LEITCH IJ, ANADON-ROSELL A, et al., 2013. Recent updates and developments to plant genome size databases [J]. Nucleic Acids Res, 42: D1159–D1166.

HU X, ZHAO EH, 2008. A preliminary study on distant hybridization of *Opisthopappus* with *Chrysanthemum* and *Ajania* [J]. Modern Agr Sci, 6: 13-14. [胡枭, 赵恩惠, 2008. 太行菊属与菊属亚菊属远缘杂交试验初报 [J]. 现代农业科学, 6: 13-14.]

HUO KS, ZHAO DL, CHEN YL, et al., 2019. Analysis of genome size and characteristics of salt-tolerant plant *Ipomoea pes-caprae* (L.) R. Brown [J]. J Plant Genet Resourc, 20(3): 728-735. [霍恺森, 赵冬兰, 陈艳丽, 等, 2019. 甘薯属耐盐植物马鞍藤基因组大小及特征分析[J]. 植物遗传资源学报, 20(3): 728-735.]

LI GQ, SONG LX, JIN CQ, et al., 2019. Genome survey and SSR analysis of *Apocynum venetum*[J]. Biosci Rep, 39(6): BSR20190146.

LI JY, LU TQ, YANG JB, et al., 2020. Genome survey and analysis of SSR molecular markers on *Macaranga indica* and *M. denticulate* [J]. Guihaia, DOI:10.11931/guihaia.gxzw202001024. [李江莹, 陆添权, 杨俊波, 等, 2020. 印度血桐与中平树基因组调查及 SSR 分子标记分析[J]. 广西植物, DOI:10.11931/guihaia.gxzw202001024.]

LIU B, SHI Y, YUAN J, et al. 2013. Estimation of genomic characteristics by analyzing k-mer frequency in de novo genome projects [J]. arXiv: Genomics, 1308.

OBERPRIELER C, HIMMELREICH S, VOGT R. 2007. A new subtribal classification of the tribe Anthemideae (Compositae) [J]. Willdenowia, 37: 89-114.

PENG Y, LAI Z, LANE T, et al., 2014. De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms [J]. Plant Physiol, 166: 1241–1254.

SCAGLIONE D, REYES-CHIN-WO S, ACQUADRO A, et al., 2016. The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F1 progeny [J]. Sci Rep, 6: 19427.

SHEN Q, ZHANG L, LIAO Z, et al., 2018. The Genome of *Artemisia annua* provides insight into the evolution of Asteraceae family and artemisinin biosynthesis [J]. Mol Plant, 11: 776-788.

SONG C, LIU Y, SONG A, et al., 2018. The *Chrysanthemum nankingense* genome provides insights into the evolution and diversification of *Chrysanthemum* flowers and medicinal traits [J]. Mol Plant, 11:1482–1491.

SONG LX, LI GQ, JIN CQ, et al., 2019. Whole genome sequencing and development of SSR markers in *Apocynum cannabinum* [J]. J Plant Genet Resour, 20(5): 1309-1316. [宋立肖，李国旗，靳长青，等, 2019. 大麻状罗布麻的全基因组分析和 SSR 标记开发[J]. 植物遗传资源学报, 20(5): 1309-1316.]

TANG Q, MA XJ, MO CM, et al., 2015. Genome survey analysis in *Siraitia grosvenorii* [J]. Guihaia, 35(6): 786-791. [唐其，马小军，莫长明，等, 2015. 罗汉果全基因组 Survey 分析[J]. 广西植物, 35(6): 786-791.]

VINOGRADOY AE. 2003. Selfish DNA is maladaptive: evidence from the plant Red List [J]. Trends Genet, 19(11): 609-614.

VURTURE GW, SEDLAZECK FJ, NATTESTAD M, et al. 2017. GenomeScope: fast reference-free genome profiling from short reads [J]. Bioinformatics, 33(14): 2202-2204.

WANG CR, YAN HD, LI J, et al., 2018. Genome survey sequencing of purple elephant grass (*Pennisetum purpureum* Schum 'Zise') and identification of its SSR markers [J]. Mol Breed, 38(7): 94-104.

WANG X, ZHOU JY, SUN HG, et al., 2018. Genomic survey sequencing and estimation of genome size of *Ammopiptanthus mongolicus* [J]. J Plant Genet Resourc, 19(1) : 143-149. [王雪，周佳熠，孙会改，等, 2018. 新疆沙冬青基因组调查测序与基因组大小预测[J]. 植物遗传资源学报, 19(1) : 143-149.]

WU YF, XIAO FM, XU HN, et al., 2014. Genome survey in *Cinnamomum camphora* L. Presl [J]. J Plant Genet Resourc, 15(1): 149-152. [伍艳芳，肖复明，徐海宁，等, 2014. 樟树全基因组调查[J]. 植物遗传资源学报, 15(1): 149-152.]

YE H, WANG Z, HOU H, et al., 2021. Localized environmental heterogeneity drives the population differentiation of two endangered and endemic *Opisthopappus* Shih species [J]. BMC Ecol Evol, 2021, 56.

YU FL, HUANG M, ZHANG HB, et al., 2019. Genome survey and characteristic analysis of SSR in *Callicarpa nudiflora* [J]. China J Chin Mat Med, 44(18): 3974-3978. [于福来，黄梅，张影波，等, 2019. 裸花紫珠基因组调研及 SSR 特征分析[J]. 中国中药杂志, 44(18): 3974-3978.]

ZHAO HB, CHEN FD, CHEN SM, et al., 2010. Molecular phylogeny of *Chrysanthemum*, *Ajania* and its allies (Anthemideae, Asteraceae) as inferred from nuclear ribosomal ITS and chloroplast *trn* L-F IGS sequences [J]. Plant Syst Evol, 284(3): 153-169.

ZHAO HB, MIAO HB, WU GS, et al., 2010. Intergeneric phylogenetic relationship of *Dendranthema* (DC.) Des Moul., *Ajania* Poljakov and their allies based on amplified fragment length polymorphism [J]. Sci Agr Sin, 43(6): 1302-1313. [赵宏波，缪恒彬，吴国盛，等, 2010. 基于 AFLP 的菊属、亚菊属及其近缘属的属间关系 [J]. 中国农业科学, 43(6): 1302-1313.]

ZHAO L, ZHU YH, WANG M, et al., 2021. Estimation of *Rehmannia glutinosa* genome size based on flow cytometry and genome survey analysis [J]. Chin Trad Herb Drugs, 52(3): 821-826. [赵乐，朱畇昊，王敏，等, 2021. 基于流式细胞术和基因组 survey 分析的地黄基因组研究[J]. 中草药, 52(3): 821-826.]

ZHEGN Y, HE ZK, YAO MJ, et al., 2020. Genome survey study of *Alpinia katsumadai* based on Illumina high throughput sequencing [J]. Chin Trad Herb Drugs, 51(13): 3530-3534. [郑燕，何志凯，姚梦鹃，等, 2020. 基于 Illumina 高通量测序技术的草豆蔻基因组研究[J]. 中草药, 51(13): 3530-3534.]

ZHOU JY, WANG X, GAO F, et al., 2017. Genome Survey and SSR Analysis of *Ammopiptanthus mongolicus* [J]. Genom Appl Biol, 36(10):4334-4338. [周佳熠，王雪，高飞，等，2017. 蒙古沙冬青基因组调查及 SSR 分析[J]. 基因组学与应用生物学, 36(10):4334-4338.]